We thank Referee 2 for the careful and thorough review of our manuscript and for the valuable comments and suggestions. In the following, we provide a point-by-point response to each comment raised by the reviewer.

1. perhaps add in the title that trends are (predominantly) calculated from satellite ozone retrievals

   We thought about different titles, including adding "satellite data", but came to the conclusion, that it would make the title very long without conveying more information. We prefer to do not change the title.

2. line 60 ends with two full stops.

   We removed the second full stop.

3. lines 136-138: "At 25 km, a negative trend of around 1 % per decade between April and August in the northern hemisphere (NH) and between January and July in the southern hemisphere (SH) is noted. The significance of these trends gets smaller with altitude, showing the most significant trends at 25 km." Give a possible reason/explanation for this finding.

   Figure 3 shows the estimated trend uncertainty associated with the unit conversion. The enhanced significance around 25 km may be related to stronger vertical ozone gradients, which can amplify uncertainties arising from the conversion and gridding procedures. We added the following text:
   "This is possibly because of stronger vertical ozone or temperature gradients near this altitude, which increase the uncertainties from the non-linear ozone unit conversion and vertical axis transformation."

4. Section 3.2: can you make some (general) conclusions on the relative drifts of the individual merged datasets w.r.t. the MIM? This might be interesting to link to some "deviating" trends that were found for some of the datasets (e.g. the ones containing OSIRIS) in Fig. 13.

   The datasets containing OSIRIS do have a stronger positive drift compared to the MIM, while the other four datasets have a slight negative drift. This is true for nearly all altitudes and latitudes, while Fig. 13 shows these differences mostly in the northern hemisphere. The drifts between datasets are identified to be the largest contributor to the spread in the derived trends from MLR.

5. In Figs. 4 and 5, you show the drift uncertainties for two different time periods, but you don't mention why these different time periods are considered and if the drift uncertainties are different between those two time periods. If you include it, describe it!

   We agree with the reviewer and added a short explanation why show two different drift uncertainties in Fig. 5 and how these drift uncertainties differ. The text was

changed accordingly at line 156:
"The same procedure was applied to the total ozone datasets, and results are shown for the period since 1996 (ODS peak) and 2000 in Fig. 5. We did not include the period before the ODS peak in this figure, because two of the total ozone datasets start around 1995 (GTO-ECV and GSG), leaving us with a limited selection of merged datasets and instruments before that year. Instead we show the relative drift uncertainty for two different start years (1996 and 2000), which were used in previous studies (Weber et al., 2022; Godin-Beekmann et al., 2022; Sofieva et al., 2026). Both relative drift uncertainties look very similar and show no significant differences. The drift uncertainties in the period 2000-2024 are slightly larger than in the period starting in 1996. When comparing the drift uncertainty for the total ozone datasets to the profile ozone datasets (Fig. 4), we see that the total ozone dataset drifts are overall lower and range between 0.1 and 0.5 % per decade.."

6. line 158: the highest total ozone drift uncertainties can be seen at high latitudes and in the tropics. Why is that? Again due to the higher uncertainties in the total ozone measurements there? Explain!

   We added the following text after the line 158:
   "**Higher drift uncertainties at higher latitudes are most probably due to the combination of higher uncertainties in ozone itself (higher solar zenith angles) and a larger natural variability in ozone. It is not clear why the drift uncertainties are enhanced in the tropics. Probably the low natural ozone variability masks the drifts to a lesser degree.**"

7. lines 238-245: I'm not very familiar with the concept of accumulating proxies over their dynamically active season and could not find how it is done in earlier studies (Weber et al., 2018, 2022). How is it done in practice? For example, a winter accumulated proxy is obtained by using the original Dec value as the Dec MM, the original Dec value + the original Jan value as the Jan MM, etc?

   Yes, accumulated proxies are constructed as cumulative sums over their dynamically active season. The proxy value for a given month represents the sum of the monthly proxy values from the beginning of the season up to and including that month. This reflects the cumulative influence of dynamical transport on stratospheric ozone variability (Weber et al., 2011).

8. line 266: add a right bracket after Fig. A1

   Done

9. lines 270-272: it's a good idea to run the MLR model over the entire time period available, also when only calculation trends from 2000-2024. But I was wondering if the trends would differ significantly if only the 2000-2024 period would be regressed? Have you tried this sensitivity experiment?

In order to be able to compare results from our study to Sofieva et al. (2026) and for consistency with earlier studies e.g. Godin-Beekmann et al. (2022), we do the regression over the full period, but showing only the results after 2000. The regression using the longer period is helpful for better constraining the proxy terms in the MLR. In general, the trends evaluated with the MLR using 2000-2024 are similar to post-2000 trends using the whole period, but the trend uncertainties are larger when a shorter time period is used.

10. Fig. 7: can you make the so-called thick lines thicker? They are hard to distinguish from the thin lines on my printout.

We increased the thickness of the lines a bit. From the later comments, it seems that the PDF reader might have an issue with some of our plots here, because we do use dashed and dotted lines in the different figures. We will use a different format to save the files. hopefully this will fix this and the other issues as well.

11. Fig. 8: The dataset lines are very hard to see.

Increased the thickness and alpha of the dataset lines.

12. Fig. 9, caption: the climatological thermal tropopauses are marked by full lines, not dashed lines. You might make those full lines thicker as well.

See our reply to comment 10.

13. Lines 360-361: any explanation for the smaller SH positive trends seen in GOZCARDS, which seem to get closer to the trends of the other data when the MLR is applied on deseasonalized ozone data?

We have not examined the smaller SH positive trends of GOZCARDS in detail, but we did reproduce the results of Sofieva et al. 2026, using deseasonalized ozone data. We suspect this difference comes from the seasonality that was either removed by the deseasonalization or not caught by the MLR.

14. Line 362 and caption Fig. 10: the joint-distribution uncertainties are marked by full lines (not dashed lines)

See reply to comment 10.

15. Fig. 12: can the thick lines be made thicker?

See reply to comment 10.

16. Lines 393-396 and 405-406: why are the trends from the datasets containing OSIRIS stronger than trends from other merged datasets? What’s wrong with OSIRIS? Are these datasets also drifting from the MIM (see remark on section 3.2).

The stronger positive trends found in the OSIRIS datasets are consistent with the positive relative drifts of these datasets with respect to the MIM (Section 3.2). This suggests that differences in relative drifts may contribute to the larger trend estimates.

17. Caption Fig. 13: all error bars are full lines, not dashed or dotted on my printout.

See reply to comment 10.

18. line 418: how are the five layers above the tropopause defined?

This was indeed confusing and we changed the wording and sentences in the manuscript as follows:
"The ozone profile datasets generally start at an altitude of 10 km (Table 1) which can be an issue when the tropopause is lower than 10 km, particularly at higher latitudes. For the stratospheric column calculations, data were screened out if the three lowest 1 km layers above the tropopause had missing data. In combination with the larger uncertainties of limb measurements below 20 km, this selection criterion causes larger spreads in the profile and stratospheric column trends. "
At higher latitudes, some ozone profile measurements contain missing values above the tropopause. To account for this issue, we applied a screening criterion based on the missing 1 km layers above the tropopause. Stratospheric columns were only calculated when at least 3 km above the tropopause was valid ozone data. Missing data inbetween two valid data points (if any at all) was handled by the merged datasets themselves. No interpolation was done or needed in this analysis.
Because missing values in these high-ozone layers could potentially affect the resulting column values and the derived ozone trends, we additionally compared our approach with the method described by Sofieva et al. (2026). Here we first calculate the ozone trends of the profile ozone datasets and then integrate the columns. The comparison of both layer and column ozone trends is presented in Figure A6. The results show that the two methods produce very similar trends, particularly for the Full Regression analysis, indicating that the treatment of missing values has only a minor impact on the overall trend estimates. Therefore, results obtained with both methods are provided.

19. line 473 ends with two full stops.

Removed the second full stop.

20. Although the paper is predominantly based on satellite ozone retrievals, I somewhat miss the link with trends calculated from ground-based datasets. For the total ozone trends, you could point out if and where the satellite-retrieved total ozone trends differ from the WOUDC trends. And for the stratospheric (profile) ozone trends, you might compare those with the trend estimates appearing in Jonas et al., ACP, 2026 (https://doi.org/10.5194/acp-26-8089-2026) and Mirallie et al., egusphere, 2026 (https://doi.org/10.5194/egusphere-2026-113).

This is indeed important and we included some additional text for the comparison with WOUDC and the other total ozone dataset, as well as the stratospheric ozone trends from Mirallie et al. (2026). We included a section in the discussion about the Mirallie et al. paper, but we didn't include the Jonas et al. paper, due to its focus on the Arctic which we do not consider.

"A comparison of our results, using the LOTUS setup (Fig. A7) with recent ground-based studies (Mirallie et al., 2026) shows a general agreement in the vertical structure of the ozone trends. In both studies, the largest positive trends are found in the upper stratosphere, consistent with ongoing ozone recovery. Mirallie et al. (2026) show clear negative trends between 20 to 30 km in the tropics and more mixed results in the northern hemisphere. This mostly mirrors our results although a slightly different proxy setup was used. The most notable difference is a slightly positive trend in the northern hemisphere above 20 km in our results, while Mirallie et al. show mixed results depending on the station (Figs. 20-23). This might highlight the need to look more closely at longitudinally resolved trends in a future study."